

Origin of Watson-Crick Complementarity

by

Brian K. Davis

Research Foundation of Southern California
8861 Villa La Jolla Dr. #13595
La Jolla, CA 92037

Abstract - Watson-Crick base pairs, A=U (T), G≡C, have bi-directional H-bonds. Since a double-helix can accommodate both bi- and uni-directional H-bond patterns, this restriction indicates that self-pairing, reliant on bi-directional H-bonds, by purines and their antecedents in transitional replicators, with anti-parallel strands, shaped Watson-Crick complementarity.

Key words: transitional replicators; anti-parallel strands, purine pairs; H-bond patterns,

(doi: <http://dx.doi.org/10.13140/RG.2.2.25292.72322>)

Anti-parallel poly(2'-deoxy-ribose-phosphate) strands have long been known to form the external scaffold of a DNA duplex, whose base-pairs exhibit Watson-Crick complementarity (Watson and Crick, 1953; Franklin and Gosling, 1953). Evidence is presented here that links both features of the double-helix. This furnishes additional support for occurrence of pre-RNA replication, based on 'self-recognition', in a primal era of poly(pentose-phosphate) replicators.

Pair-forming H-bonds among the nucleobases are depicted in Fig. 1. Also shown is the H-bond pattern formed between anti-parallel (C1-C5:C5-C1) molecules of the 2-keto-pentol, D-ribulose-1,5-bis-phosphate (RuBP). All three naturally occurring pairs (*left*) contain

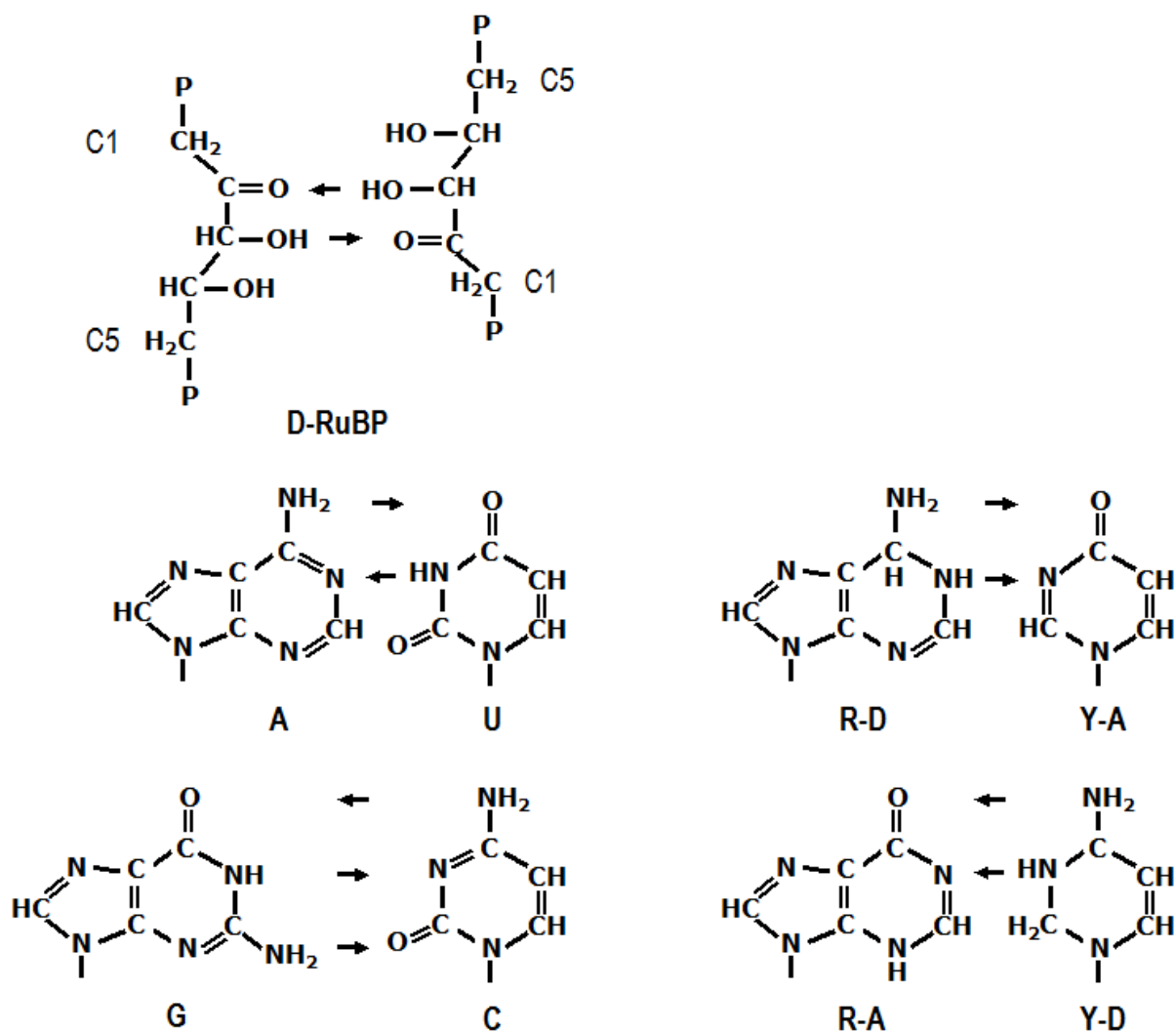


Fig. 1. H-bond patterns in pentose and nucleobase pairs. Bi-directional bonds characterize the naturally occurring pairs (*left*), while the proposed nucleobases (*right*) were designed to pair through uni-directional H-bonds. Arrows indicate donor-to-acceptor direction. RuBP, ribulose-1,5-bisphosphate; A, adenine; C, cytosine; G, guanine; U, uracil; R-D, purine H-donor; Y-A, pyrimidine H-acceptor; R-A, purine H-acceptor; Y-D, pyrimidine H-donor.

bi-directional H-bonds. Significantly, the pentose pair combines bi-directional H-bonds with an anti-parallel configuration. This achieves a form of 'self-recognition', making strand copying a feasible possibility. Watson-Crick complementarity can be seen to share a similar reliance on bi-directional H-bonds. The direct dependence of bi-directional H-bond formation, and 'self-recognition', on counter-oriented alignment, illustrated by the pentose molecules, shall be seen to provide an insight into the origin of Watson-Crick complementarity. Purines and pyrimidines designed to pair with uni-directional H-bonds (*right*), by contrast, cannot self-bond.

In addition to being an intermediate of the autocatalytic reductive pentose-phosphate cycle and precursor to ribose in the RNA-scaffold, ribulose-phosphate (Ru-P) has been identified (in its bis-phosphorylated, polymeric form) as a primal replicator (Davis, 2012, 2018). Its role as a replicator became apparent during an extensive investigation (Davis, 1999, 2006, 2008), aimed at relating regularities in genetic code structure to the evolution of amino acid synthesis pathways, prior to the Last Universal Common Ancestor (LUCA).

On embedding purine precursor (ribosyl-amine)-phosphate, (Ri-NH₂)-P, in a ladder-like poly(Ru-P, Ri-P) duplex, it could be seen to form bi-directional H-bonds with another (Ri-NH₂)-P monomer, located in the opposite (anti-parallel) strand. Synthesis of the purine intermediate on a ribose subunit (an invariant of the purine synthesis pathway), in a poly(Ru-P, Ri-P) ladder-replicator, additionally, transformed its ribose platform from a non-interacting spacer (embedded scaffold), between H-bonding Ru-P monomers, to an external scaffold, consistent with RNA structure.

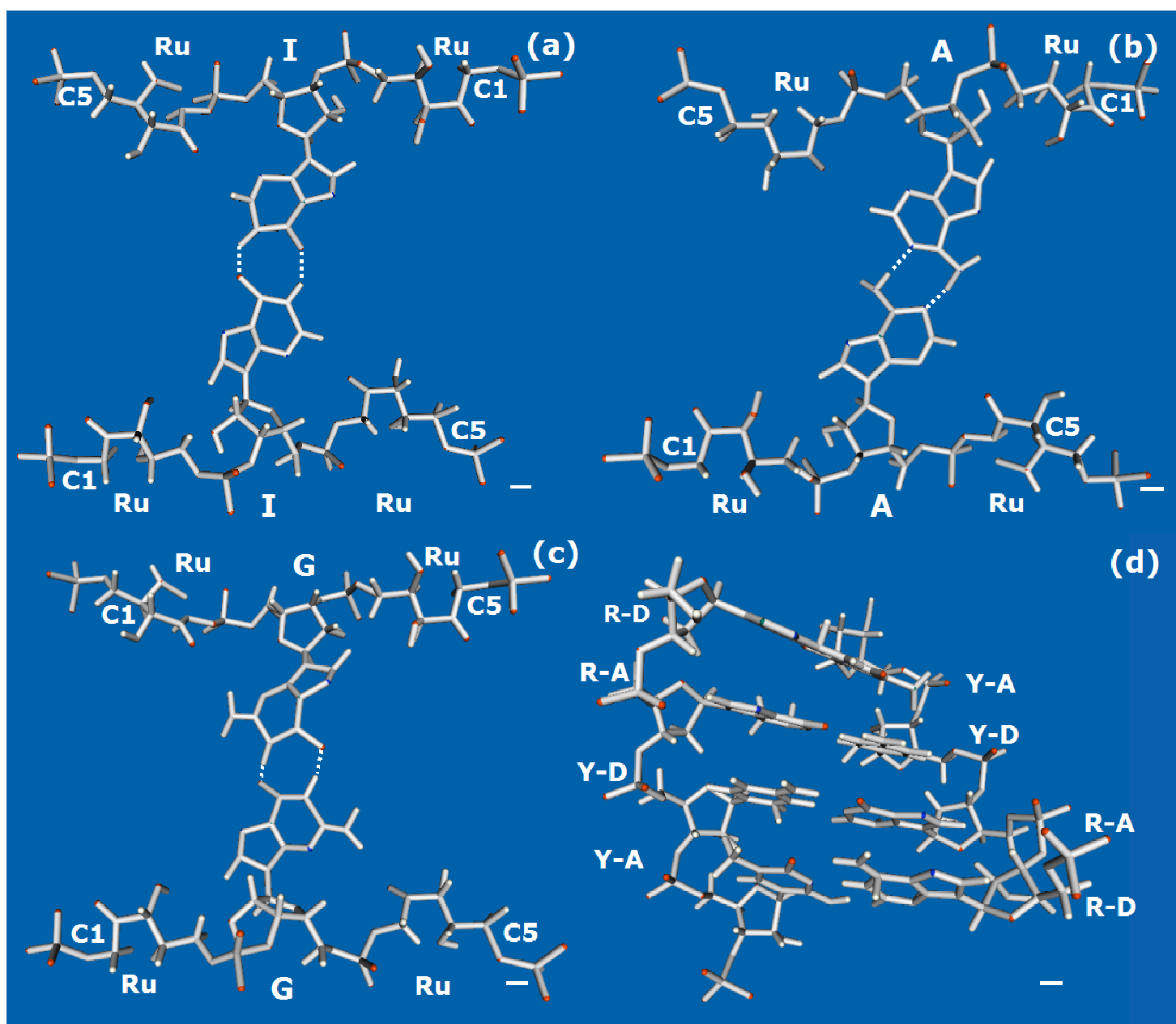


Fig. 2. Molecular models of bi- and uni-directional nucleotide H-bond interactions. (a), (b), and (c) show bi-directional H-bonds formed between designated nucleotides embedded in anti-parallel ribulose-phosphate strands. (d) An RNA double-helix (A-form) containing nucleotides designed to form non-Watson-Crick uni-directional H-bonds. A, Adenosine, G, Guanosine, I, Inosine, other labels are from Fig. 1. These molecular structures were constructed using Facio-20.1.3 model building software (Suenaga, 2005), and optimized by Gamess restricted Hartree-Fock fragment molecular orbital calculations (Schmidt et al., 1993).

Each purine synthesized in this manner could be anticipated, therefore, to self-interact, when embedded in the anti-parallel strands of a ladder-replicator, exemplified here by poly(Ru-P,Ri-P). Molecular models in Fig. 2 confirm this for Inosine, Adenosine, and Guanosine. Bi-directional H-bonds pair each self-bonded purine. It is also demonstrated (Fig. 2d) that uni-directional complementarity between purine:pyrimidine pairs are admissible within a double-helix. Since this departure from Watson-Crick complementarity is stereochemically acceptable, the reliance of RNA and DNA on bi-directional H-bonds can be attributed to history, in the sense that this derives from earlier, ladder-replicators, with a requirement for self-interaction between anti-parallel strands.

The greater role of purines in metabolism implies they predated pyrimidines (Wächtershäuser, 1988,1992). Consistent with this, the purine synthesis pathway has a single, simple invariant, Ri-P (pre-LUCA attachment site for a cofactor, or scaffold). By contrast, the pyrimidine pathway has an initial carboxyl invariant, which is jettisoned for an Ri-P invariant, in an apparent cofactor exchange, leading to the synthesis of orotidine-phosphate.(Davis, 2015). With the synthesis of pyrimidines, equipped to pair with pre-evolved purines, double-helix formation could take place. The linear rise/base-pair decreased from 8.4 to 2.6 Å, in a ladder to double-helix transition. A substantial compression of genetic information thereby accompanied replacement of replication based on purine 'self-recognition' to one with nucleobase 'complement-recognition'.

Linking the bi-directional H-bond patterns underlying Watson-Crick complementarity to the imprint of nucleobase synthesis, preceding the double-helix, furnishes additional evidence that the poly(pentose) scaffold of RNA and DNA, itself, derived from an earlier replicator.

References

- Davis BK 1999. Evolution of the genetic code. *Prog. Biophys. Mol. Biol.* **72**, 157-243.
- Davis BK 2006. Making sense of the genetic code with the path-distance model. In *Messenger RNA Research Communications* Ed. MH Ostrovskiy, New York: Nova Science pp. 1-32.
- Davis BK 2008. Imprint of early tRNA diversification on the genetic code: Domains of contiguous codons read by related adaptors for sibling amino acids. In *Messenger RNA Research Perspectives* Ed. T. Takeyama, New York: Nova Science, pp 35-79.
doi: <http://dx.doi.org/10.13140/RG.2.2.35163.87841>
- Davis BK 2012. Replicative-form of poly(triose-phosphate).
<http://www.archive.org/ReplicativeFormOfPolyTriose-phosphate>
- Davis BK 2015. Path invariants and the evolution of replication.
<http://www.archive.org/details/EvolnOfRepln2015>
doi: <http://dx.doi.org/10.13140/RG.2.2.10002.96963>
- Davis BK 2018. Deep-structure of the genetic code and the origin of replication: path Invariants as pre-LUCA attachment sites. FEBS2018 P19-016-Mon.
doi: [10.13140/RG.2.2.20657.68967](https://doi.org/10.13140/RG.2.2.20657.68967)
- Franklin RE, Gosling RG 1953. Molecular configuration for sodium thymonucleate. *Nature* **171**, 740-741.
- Schmidt MN, Baldrige KK, Bohtz JA, Elbert ST, Gordon MS, Jensen JA, Koscki S, Matsunaga JA, 1993. Gamess general atomic and molecular electronic structure System. *J. Compu. Chem.* **14**, 1347-1363.
doi: <http://dx.doi.org/10.1002/jcc.540141112>
- Suenaga M 2005. Facio: New computational chemistry environment for PC Gamess. *J. Comput. Chem. Japan* **4**, 25-32 doi: <http://dx.doi.org/10.2977/jccj.4.25>
- Wächtershäuser G 1988. An all-purpose precursor of nucleic acids. *Proc. Natl. Acad. Sci.* **85**, 1134-1138
- Wächtershäuser G 1992. Groundworks for an evolutionary biochemistry: the iron-sulfur world. *Prog. Biophys. Mol. Biol.* **58** 85-201.
doi: [http://dx.doi.org/10.1016/0079.6107\(92\)90022](http://dx.doi.org/10.1016/0079.6107(92)90022)
- Watson JD, Crick FHC 1953, A structure for deoxy-ribose nucleic acid. *Nature* **171**, 737-738